ON THE VARIABILITY OF INDIVIDUAL JUDGMENTS

BY

FREDERICK LYMAN WELLS, PH.D.

Assistant in Pathological Psychology, McLean Hospital, Waverley, Massachusetts.

[Reprinted from *Essays Philosophical and Psychological in Honor of William James.* New York, 1908]

ON THE VARIABILITY OF INDIVIDUAL JUDGMENTS

By Frederic Lyman Wells

In the article "Statistics of American Psychologists" [1] Professor Cattell calls attention to the fact that if one endeavors to arrange and rearrange in serial order a number of given objects, the positions successively given them will vary somewhat as they would vary if the arrangements had been made one each by different observers. If we undertook to rearrange ten times a series of grays in order of brightness, we should no more get the same order each time than we should get identical orders from ten different subjects. Nor would our own orders vary approximately the same amount from the average; sometimes we should be better, sometimes worse, judges, just as among our ten subjects some would be more discriminative, some less. The judgments of the same individual at different times are theoretically quite comparable to those of different individuals regardless of the factor of time.

In this way there may be illustrated a continuum between the subjective and objective classes

[1] *Am. J. Psych.*, Vol. XIV, 320–328.

of judgment. In the case of grays, weights, or lines we assume a certain standard which we term the objective order, and which we determine through photometry or some analogous method. Because we have such methods, we do not need to have recourse to individual judgments to determine objective values, and these individual judgments give us a part of the personal equation; the individual's sensibility to light, weight, etc. On the other hand, we have such subjective judgment as preferences in sculpture, painting, or music. In the first class we may arrange individuals in precise order for accuracy of discrimination; in the second, one may with equally good taste vary his preferences within a considerable range. So far as any distinction on a statistical basis is possible, we might consider as subjective those types in which the various judgments of the individual formed a species of their own, varying from each other considerably less than from an equal number of judgments made by different individuals; and consider as objective those in which an individual would vary from his own independent judgments about as much as the variation of an equal number of judgments by different individuals. For example, if A and B arranged ten pieces of music in order of preference, the orders would centre about each individual's own standard; but if A, B, C, D, etc., arranged ten graduated weights, the orders would theoretically all centre about a

common standard, the objective order of heaviness. The two categories would almost certainly be continuous. We may first consider from this viewpoint types of this first, or highly subjective, class of judgments, and compare these subsequently with examples of a more objective type.

Experiments in Preference

An obvious and serious difficulty with all experiments involving repeated judgments of the same thing are the factors of recognition and memory. Especially is this true of judgments of subjective preference with which we are to be here concerned. If the subject remembers his previous judgments, he will in spite of himself order his successive ones accordingly. The only practicable ways of meeting this difficulty are to make the series to be arranged as long as possible, and to allow as much time as possible to elapse between the successive arrangements. A certain homogeneity in the series is necessary, and this made the selection of suitable material no easy task. A series of fifty colored souvenir postal cards, to be graded in order of individual preference, was finally decided upon as the most practical approach to the problem.[1] The cards were approved

[1] The psychological possibilities of the souvenir postal card have been insufficiently appreciated. They afford an inexhaustible mine of material for experiments in recognition memory and kindred processes, for which there is no other readily accessible apparatus.

by the writer from selections made from the sample books of the Rotograph Company. They are all views of natural scenery, with the works of man a subordinate feature. In a few cards these last are altogether absent. The fifty cards were arranged by the five subjects, A–E, five times each, one week elapsing between each individual's successive arrangements. Single arrangements were also made by five additional subjects, F–J, and these, combined with the first arrangements of A–E, give, for comparison, a series of ten arrangements by different subjects. Subjects A, B, and C are men of special psychological training, D and E are women of moderate psychological training. Of the five subjects making single arrangements, all are men of special, though widely differing, psychological training. From these experiments are gathered the data to be discussed below.

The uniform attitude of the subjects toward the experiment was one of lack of confidence in the judgments. The time required to make a single arrangement varied from 15 to 45 minutes, the women taking as a rule longer than the men, and the time, of course, decreasing with the successive arrangements. So far as exact positions of the cards were concerned, the subjects who made repeated judgments reported complete oblivescence except now and then with regard to first or last positions. Of course a remembered judgment

was not necessarily repeated nor were repeated judgments necessarily remembered; subject E placed the same card last in each arrangement, and at the close expressed surprise at finding that she had done so. One subject expressed absolute certainty that new cards were being successively introduced. There was naturally subjective effort to judge independently of previous arrangements. Certain features are to be noted in the results indicating that the memory difficulty was fairly satisfactorily met.

The subjoined Table I gives under X the order, average positions, and m. v. (not p. e.) of the single arrangements by the ten subjects. Column V is a combination of the records of subjects A–E which will be described below. Table II gives in detail the results of the five successive arrangements by each of the subjects A–E. To anyone interested in the statistics of such arrangements they will perhaps repay a more careful examination than it is possible to give them here.

When the subjects made the arrangements, it was customary to hesitate considerably on the first few and then to proceed at about an equal, or perhaps slightly increasing, rate to the end. This hardly reflects the size of the differences, which are presumably greatest at the ends. It is due merely to a natural tendency to exercise greater care at the beginning of the experiment. So far as the actual order is concerned they cannot have

TABLE I

X (Results for Ten Subjects)				V (Av. of Table II)		
Order.	Roto-graph Co. Serial No.	Position.	M. V.	Order.	Position.	M. V.
1	5442	12.6	10.2	1	2.6	1.3
2	5511	13.3	8.8	2	4.0	2.3
3	5353	13.6	7.7	3	5.8	3.1
4	2460	14.1	6.3	4	6.8	2.9
5	7384	15.0	10.4	5	7.4	3.8
6	106 *b*	15.8	9.6	6	8.4	4.4
7	30 *a*	16.7	9.6	7	9.6	6.4
8	6151	17.0	10.8	8	10.2	2.3
9	8708	17.4	7.0	9	10.6	2.3
10	5521	17.6	4.6	10	12.0	5.3
11	7118	18.0	11.1	11	12.8	4.5
12	7198	19.4	10.2	12	13.4	5.2
13	6236	20.6	8.8	13	14.0	5.1
14	7196	21.1	14.7	14	15.0	5.0
15	3893	21.4	13.2	15	15.8	7.0
16	2012	22.0	9.2	16	17.6	6.0
17	6182	22.5	10.7	17	18.6	5.0
18	5626	22.7	10.7	18	19.2	5.1
19	7570	23.0	11.4	19	19.8	6.2
20	6976	23.4	10.4	20	21.0	6.0
21	6156	23.4	10.8	21	21.6	6.4
22	5560	23.9	12.9	22	21.8	4.9
23	7125	24.4	10.6	23	23.0	5.4
24	5710	24.5	10.3	24	25.0	5.6
25	7171	25.1	16.4	25	26.0	7.0

TABLE I — *continued*

X (Results for Ten Subjects)				V (Av. of Table II)		
Order.	Roto-graph Co. Serial No.	Position.	M. V.	Order.	Position.	M. V.
26	5871	26.4	11.1	26	26.8	4.7
27	911	26.5	14.1	27	27.4	6.0
28	7522	27.0	12.2	28	28.0	6.0
29	184	27.0	16.4	29	28.6	3.9
30	16103	27.4	10.8	30	29.2	4.3
31	6264	27.6	8.0	31	30.0	6.5
32	7170	28.1	6.7	32	30.4	6.2
33	5731	28.8	8.8	33	32.2	5.0
34	5439	29.1	12.1	34	33.0	4.8
35	7197	29.3	9.3	35	33.6	5.8
36	8706	29.6	8.2	36	34.4	6.8
37	5570	29.9	14.5	37	34.6	5.0
38	25508	30.2	11.4	38	35.0	5.0
39	6442	30.3	10.9	39	36.2	4.8
40	6547	30.4	9.2	40	36.4	6.0
41	5727	30.5	16.7	41	37.4	4.8
42	8704	31.0	11.0	42	39.2	4.0
43	2103	32.6	12.3	43	40.4	5.8
44	6670	32.7	13.0	44	42.0	3.5
45	6976 *a*	34.8	5.9	45	43.2	3.9
46	7026	35.5	8.1	46	44.6	3.0
47	2010	36.6	8.6	47	45.4	2.0
48	5862	36.6	10.8	48	46.6	2.0
49	5860	38.4	10.2	49	48.0	1.2
50	1285	43.1	5.1	50	49.6	0.6

TABLE II

Order	Rotograph Co. Serial No.	A		Rotograph Co. Serial No.	B		Rotograph Co. Serial No.
		Position	M. V.		Position	M. V.	
1	5353	1	0.0	7384	4	1.8	6156
2	5521	3	1.2	5442	6	4.4	6264
3	106 *b*	4	2.0	2460	7	5.2	5731
4	5560	6	1.8	25508	8	4.8	6670
5	5511	7	3.4	6151	9	3.0	8704
6	6182	7	3.8	2012	10	7.6	2102
7	6151	8	0.8	7196	11	9.6	7196
8	5442	8	2.0	5560	12	2.6	5511
9	6976	9	3.4	5353	13	8.4	7197
10	30 *a*	12	4.4	8708	14	9.4	7198
11	5871	13	3.6	7522	15	6.2	16103
12	7125	14	3.8	106 *b*	15	9.8	8708
13	7118	15	4.8	6442	15	8.8	5521
14	5710	15	0.8	6226	18	7.6	5710
15	7384	15	4.6	5439	18	11.2	5727
16	2460	17	6.4	8706	19	9.4	5570
17	7522	17	3.4	6156	19	3.4	184
18	911	18	2.4	5871	19	6.2	7384
19	5626	18	2.2	2013	19	10.0	5303
20	6226	19	4.2	6182	20	6.0	8706
21	7570	20	1.8	911	20	12.2	6151
22	16103	20	5.4	3893	21	6.2	5442
23	6156	23	6.0	309	21	12.2	911
24	5731	26	2.8	5710	25	4.0	2013
25	1285	26	6.4	6796	26	9.0	106 *b*

TABLE II — *continued*

C		Roto-graph Co. Serial No.	D		Roto-graph Co. Serial No.	E	
Posi-tion	M. V.		Posi-tion	M. V.		Posi-tion	M. V.
3	0.8	7384	3	3.0	7384	2	1.0
4	3.4	5353	3	1.2	5511	4	1.2
6	1.4	6976	6	3.4	5442	6	3.4
7	3.2	6151	6	1.6	16103	7	3.0
7	2.8	5521	7	4.8	5353	7	4.8
9	4.2	3893	8	3.4	2460	8	3.1
11	3.8	5511	9	5.2	5521	9	5.6
13	8.0	5710	9	5.6	5676	9	4.2
13	6.8	5560	9	1.6	3893	9	2.2
13	2.4	6182	10	2.0	6264	11	8.4
14	4.8	7125	10	2.6	6226	12	5.4
15	6.0	2460	11	2.6	7118	12	3.6
15	4.0	106 *b*	11	3.0	5727	14	4.8
15	8.2	30 *a*	11	5.2	5439	16	3.4
16	10.2	7118	12	3.6	6976	18	5.6
18	6.4	5442	15	3.4	7198	19	4.6
19	3.2	6670	17	3.2	184	21	11.8
20	8.6	7570	18	2.6	6442	21	5.8
22	8.0	8708	19	2.8	7125	21	7.8
22	9.6	6442	22	3.4	106 *b*	22	6.6
23	9.6	911	23	3.8	6976 *a*	22	4.8
23	3.0	6264	23	2.2	7570	22	7.6
23	6.8	6226	24	4.0	6547	24	12.8
24	9.0	7522	25	5.0	5862	25	7.2
25	6.8	7196	26	6.2	5731	27	6.6

TABLE II — *continued*

Order	Rotograph Co. Serial No.	A		Rotograph Co. Serial No.	B		Rotograph Co. Serial No.
		Position	M. V.		Position	M. V.	
26	7198	27	6.2	7125	26	8.4	6976
27	3893	28	6.8	2010	26	6.4	6182
28	5570	28	3.0	5862	27	10.4	6226
29	7196	29	4.0	718	28	5.6	30 *a*
30	7170	30	2.0	7026	29	5.6	7171
31	184	31	8.6	5321	29	4.2	7170
32	6264	32	3.2	1285	29	9.2	5871
33	6547	34	3.4	7198	31	8.6	5560
34	7197	34	2.4	8704	32	5.8	7570
35	5439	35	6.8	6264	32	6.0	6547
36	7171	36	4.2	7170	32	7.8	2010
37	8704	36	3.2	5511	33	3.0	2460
38	8708	36	2.8	5626	34	5.2	5626
39	2012	37	5.2	16103	35	4.2	7522
40	5727	37	4.8	6976 *a*	35	8.6	3983
41	6976 *a*	38	5.6	5370	35	8.2	25508
42	6670	40	3.6	6547	38	4.4	5439
43	8706	41	4.4	5727	38	8.8	1285
44	25508	44	1.4	5731	41	6.0	7125
45	6442	45	2.2	7170	42	3.4	7118
46	5862	45	0.8	5860	44	3.6	6976 *a*
47	2013	46	2.0	6670	45	1.0	7026
48	7026	47	1.2	184	48	1.2	6442
49	2010	48	1.2	7171	49	1.0	5862
50	5860	50	0.0	7197	49	1.6	5860

The positions are given to the nearest positive integer only; the orders

TABLE II — *concluded*

C		Roto-graph Co. Serial No.	D		Roto-graph Co. Serial No.	E	
Posi-tion	M. V.		Posi-tion	M. V.		Posi-tion	M. V.
27	3.0	25508	27	3.6	2012	27	2.2
28	5.2	5439	27	4.6	30 *a*	28	7.0
28	4.2	5871	28	4.8	5710	29	7.6
29	5.6	5626	28	2.2	7197	29	2.2
29	4.8	6156	28	4.2	7170	30	5.0
30	8.8	7198	29	2.4	8708	31	8.6
30	5.6	7170	29	4.2	7196	32	9.0
31	7.0	6976 *a*	32	2.6	6182	33	3.2
32	6.4	16103	34	1.4	6151	33	8.2
32	9.8	2012	36	2.0	8706	33	4.6
33	8.4	7197	37	5.6	5860	34	8.0
33	8.8	5862	37	2.4	6670	34	7.4
33	9.4	5860	37	1.8	7026	35	5.6
33	9.6	2013	40	1.6	5871	36	3.6
33	10.6	2010	40	1.4	25508	37	4.4
36	3.4	5727	41	2.8	5560	37	3.8
39	4.6	7026	42	2.4	7171	37	5.2
40	7.2	5731	43	2.6	911	40	5.8
41	2.0	5570	44	2.6	2010	40	5.6
42	5.2	6547	45	0.6	2013	42	8.2
45	3.0	8704	46	3.2	8704	43	4.4
47	0.8	7171	46	2.6	7522	43	3.6
47	3.0	8706	47	1.6	6156	44	3.2
47	2.2	184	48	0.8	5570	48	0.8
50	0.0	1285	49	1.6	1285	50	0.0

are correct to a smaller scale, equal positions being indicated by brackets.

much bearing upon the experimental study of æsthetics, because the material would be too difficult to standardize for this purpose. Certain of the cards necessarily fall into groups through similarity of subject or color scheme, and these tend to keep rather together in position, also through the fact that they tend to become associated in memory. So far as establishing any objective basis for criteria of preferability is concerned, the results seem to me almost entirely negative.

It will perhaps be easier to consider in some detail the figures in Table I as a preliminary to the special results of the repeated arrangements in Table II. Column X presents almost a chaos of variability, the extreme range barely covering 30 places, with one exception only 26. The m. v.'s average nearly 11 places and range from the least variable card with an m. v. of 4.6 to the most variable with an m. v. of 16.7 over an approximately normal distribution as follows:

Variation	5	6	7	8	9	10	11	12	13	14	15	16	17
No. cases	2	2	1	5	7	8	13	3	3	2	1	2	1

Among the individual variations there are many above 25, the highest being 32. Card 2460, in which this variation occurs, has an average position of 33, and the individual places assigned to it by the ten subjects are respectively 42, 1, 40, 37, 43, 2, 42, 42, 41. A card graded first by one subject was in two cases graded last by another; in a third, next to the last. One of the former is the

most variable card, 5727, and its grades are respectively 48, 44, 1, 43, 6, 48, 24, 33, 8, 50. The grades of the least variable card, 5521, are 2, 32, 17, 14, 20, 20, 15, 18, 20, 18; position 17.6. Anyone acquainted with the meaning of such figures as those given above must recognize the futility of attempting to evolve from them an order of any objective value.

This is much modified in the repeated arrangements by the same subject. It was noted above, that in objective judgments, as of weights, we should, theoretically, vary as much from ourselves as other people varied from each other, and from the comparison of these two variabilities might be deduced the degree of objectivity of the judgments. In the repeated arrangements it is at once evident that the range is much greater and the variability smaller. A table most comparable to X is given under V, which is computed as follows: Subject A's best card, as will be seen from Table II, receives an average of 1, B's an average of 4, C's 3, D's 3, and E's 2. Thus the average position of the best card of the five repeated judgments by five subjects is 2.6, and the average of the respective m. v.'s is 1.3, as opposed to 12.6 and 10.2 for individual judgments by ten subjects. The figures for last position are seen to be 49.6 and .6 as against 43.1 and 5.1. Of course the extreme positions might unduly favor the repeated judgments in this respect. But the figures for the

middle five judgments are respectively 24.6 and 5.7 as against 25.6 and 12.4. Table III below gives a basis for a more complete comparison of the two variabilities. Each series in Table II contains 50 average judgments, consequently 50 m. v.'s in all. These have been divided into 10 consecutive groups of 5 each. Thus under 1–5 and opposite A we find 1.7, which is the average of the m. v.'s of the five cards which stood highest as a result of A's five consecutive arrangements. Under 15–20 and opposite D is 3.1, the average m. v. of cards 16–20 from the series of five arrangements by D, etc. Opposite Av. are given the averages of the five subjects for each set of five consecutive positions. At the bottom are given the average m. v.'s for the various groups of positions as assigned by the ten subjects.

TABLE III

AVERAGE M. V. FOR EACH SET OF FIVE CONSECUTIVE POSITIONS

Subject	Positions									
	1–5	5–10	10–15	15–20	20–25	25–30	30–35	35–40	40–45	45–50
A	1.7	2.9	3.5	3.7	4.5	4.4	4.9	4.0	3.4	1.0
B	3.8	7.5	8.7	7.0	8.7	7.3	6.8	5.8	6.2	1.7
C	2.3	5.0	6.6	7.2	7.0	4.6	7.5	9.8	4.5	1.8
D	2.8	3.6	3.4	3.1	4.2	3.9	2.5	2.6	2.2	1.9
E	2.7	4.7	4.6	7.3	7.8	4.8	6.7	4.8	5.7	2.4
Av.	2.6	4.7	5.4	5.7	6.4	5.0	5.7	5.4	4.4	1.8
Ten subjects	8.7	8.3	11.6	10.5	12.2	12.9	10	10.8	11.8	8.5

In examining the portion of this table dealing with the repeated arrangements we find, as we should anticipate, that the m. v. increases toward the middle positions and decreases toward the ends. The amount of this increase varies considerably, and constitutes a not uninteresting point of individual difference. In subject A the middle m. v.'s are nearly three times those at the start; in D they are barely half again as much. Individual difference in reliability of judgment seems therefore to be greater in the middle than at the ends. This is what we should expect, for the judgments are more difficult in the middle, and we naturally vary more from each other in our judgment of difficult things than in our judgment of easy ones. Another point of significance is that the m. v.'s are always less at the disliked than at the preferred end, although there is no intrinsic reason why they should be better grounded in memory. This might be in part due to a generally unæsthetic series of cards, but it is perhaps generally true that we are surer of our antipathies than of our preferences.

In the m. v.'s of the ten subjects the most striking appearance beyond their greater size is that the increase in the middle and the decrease at the ends is not nearly so well marked as in the repeated arrangements. This is precisely the condition that the memory factor in the repeated arrangements would give, but under Table IV will be cited

reasons against its being due to this cause. It is also to be noted that here the m. v.'s of the disliked end are not smaller than those of the preferred, though the difference is insignificant.

The m. v.'s of the repeated arrangements of subjects A–E are shown according to series in Table IV *a*.

TABLE IV *a*

COMPARATIVE VARIABILITY OF THE INDIVIDUAL SERIES

Subject	Series I	II	III	IV	V	Av.
A	4.95	3.04	3.08	2.98	3.08	3.43
B	6.82	6.06	4.82	6.30	7.54	6.21
C	6.64	4.84	5.00	6.88	4.72	5.61
D	3.88	3.26	2.58	2.76	2.80	3.06
E	6.72	4.78	5.06	4.48	4.62	5.13

Thus under I–A we find 4.95, which is the average variation of the judgments made in A's first arrangement from the average of the five arrangements made by him; 3.04 is the variation of his second arrangement, etc. Through this table we can determine what arrangement, if any, tends to be the most accurate. In subject A the fourth is the most accurate (av. m. v. 2.98), in subject B the third, C the fifth, D the third, E the fourth. Now assuming any considerable operation of the memory factor in these experiments, one of two

things should result. Either the first judgment should set the standard from which the successive arrangements would vary more or less, or, as the memory of previous judgment accumulates, each successive judgment would become more and more the sum of the preceding arrangements, and the m. v. would progressively decrease. The latter event seems to the writer the more likely, but neither is recorded in the figures, save in so far as the first judgment tends to be a relatively inaccurate one. It is difficult to see in what way the relative accuracy of the successive judgments is distributed differently from what it might be if the successive arrangements had been made by different individuals. They seem to be quite as independent of one another.

Whatever the effect of the memory factor upon the successive series of judgments, those at the ends should be most susceptible to it, those in the middle least. The proper procedure is, then, to examine the variability in the succeeding series according to the position of the cards, and to note if there is any difference in the variability of the successive series according as the positions are high, intermediate, or low. Table IV *b* gives for each subject the variability of the first five, the middle five, and the last five positions, in each of the successive series. No significant difference appears in the relative size of the variabilities of the middle and end cards, *according as the suc-*

TABLE IV *b*

	I	II	III	IV	V
Position 1–5					
A	1.8	1.0	2.2	1.0	2.4
B	3.4	3.4	3.6	5.2	3.6
C	2.2	1.8	1.2	2.8	2.6
D	5.0	2.0	2.4	2.0	2.0
E	2.4	1.4	3.2	4.2	2.4
Av.	3.0	1.9	2.7	3.1	2.6
Position 23–27					
A	7.8	4.8	6.8	3.6	4.2
B	8.8	11.6	6.2	4.0	6.8
C	8.8	4.6	7.4	7.4	4.6
D	4.4	4.0	4.0	5.4	5.2
E	11.0	5.2	6.6	6.4	10.6
Av.	8.2	6.0	6.2	5.4	6.3
Position 46–50					
A	1.4	0.8	1.6	0.6	0.8
B	2.2	1.4	1.4	1.8	1.6
C	1.0	1.6	1.6	0.6	1.6
D	2.8	2.4	1.6	1.0	2.2
E	2.4	5.0	1.8	0.8	2.0
Av.	2.0	1.3	1.6	1.0	1.6

cessive series are reached. If memory has operated at all, it must have operated in positions 1–5 and 46–50; from positions 23–27 it is practically excluded. As there is nothing save consistent differences in size to distinguish them, it seems justifiable to infer that memory has in no way made the end judgments less independent than the middle ones. For this reason also, some other explanation must be assigned to the fact that the m. v.'s of the middle and end positions in the repeated arrangements are more different than those of the analogous positions in the individual arrangements by the ten subjects.

In the last column of Table IV *a* are given the averages of the m. v.'s of each series, the total variability of the five successive series for each subject. There is here a difference of about 2:1, B varying the most from his own judgments with 6.21, D the least with 3.06. The average of all the variabilities is 4.7. Following are the variations of each of the ten subjects from their average:

TABLE V

A	B	C	D	E	F	G	H	I	J	Av.
9.34	10.94	12.98	8.68	11.54	10.34	12.46	9.32	9.12	9.34	10.48

A somewhat significant comparison is afforded between the variability of subjects A–E from the average of the ten, and their variation from their own judgments as given in Table IV *a*. Those who vary least from their own judgments also

vary least from the judgments of others. Thus D, whose preferences are the most consistent with her own, also agrees best with the judgment of others. A is next in both (among subjects A–E), and the entire orders agree with 20 per cent of displacement. The observations are too few to do more than suggest a general principle, but their interpretation is a rather interesting one. The critic who best knows his own mind would seem the best criterion of the judgments of others. I have elsewhere argued, mainly on theoretical grounds, against the validity of accepting the accordance of a judgment as indicative of its accuracy, but figures like the above are an empirical demonstration in its favor. This matter will be recurred to towards the close of this paper.

With respect to such judgments as those with which we are dealing, the variability of different individuals is seen to be more than twice as great as the variability of different judgments by the same individual. Each individual's judgments form a distinct species of their own, and the opinions expressed are thus in a high degree personal and subjective.

Brief attention may be called to the character of the individual variations themselves. The distribution of the m. v.'s for the averages of the ten subjects has already been given. For the five consecutive judgments of subjects A–E, the m. v.'s are distributed as follows:

TABLE VI

DISTRIBUTION OF THE MEAN VARIATIONS OF EACH SUBJECT

Subject	Variation	0	.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10	10.5	11	11.5	12	12.5
A	No. cases	2	3	4	2	8	2	5	6	5	3	2	1	4	2	..	..	..	1	..	..	..	..	..	..	..	..
B	"	..	..	3	2	..	1	4	1	5	1	2	3	7	..	..	3	3	4	4	2	2	..	1	..	2	..
C	"	1	2	2	..	4	1	6	1	3	3	2	2	2	3	2	..	4	3	1	5	2	1	..	..	..	..
D	"	..	2	3	6	7	9	8	3	3	3	3	2	1	..	..	..	..	..	..	..	..	..	..	..	..	..
E	"	1	1	2	..	3	..	6	5	4	5	3	6	..	2	3	2	3	1	1	..	..	..	..	1	..	1

There is a suggestion of species in the distributions for subjects B and C, as though there were a type of card in which the judgments were likely to vary more than in others. The remainder do not show this characteristic. The largest single mean variation is 12.8, made by subject E on card 6264, which stands 31st in this subject's series. The zero cases are from first and last places, with one exception presumably remembered from time to time.

Following are the distributions of the individual variations in the successive judgments. They are

TABLE VII

DISTRIBUTION OF THE SINGLE VARIATIONS

Subject	Variation	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
A	No cases.	28	56	43	30	23	20	19	5	9	1	..	2	..	..	1	1	1	..	..	..	..	1	..	..	..	..	..	..	..	..
B	"	13	32	27	20	18	23	15	22	14	12	12	9	5	7	3	5	4	1	3	2	1	2	1	2	..	..	..	..	..
C	"	14	28	35	16	18	26	17	12	10	15	10	7	8	6	5	4	3	1	3	1	..	..	..	..	..	..	..	..	..
D	"	26	47	47	36	23	18	21	12	5	3	2	1	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
E	"	21	31	22	27	30	16	23	18	10	10	4	6	4	3	2	4	2	3	1	1	1	2	..	..	..	..	..	..	..

ordinary skew distributions with no striking features. The variability of the single judgment seems

to be distributed practically according to chance, limited, of course, at the small end.

In a previous study [1] attention was called to the fact that in many consecutive orders the difference in position as indicated by the average did not bear a very strict relation to the reliability of the judgments as given in the probable error. Small differences might exist side by side with small p. e.'s, and large differences with large p. e.'s. On account of the lack of material for empirical analysis the question was merely indicated, but an examination of the longer ranges obtained in the present experiment indicates that the difference between any two consecutive positions is not given in the averages and p. e.'s or even in the entire distributions, but that some refinement of the treatment is necessary.[2]

[1] "A Statistical Study of Literary Merit," *Archives of Psychology*, No. 7, pp. 17–19.

[2] The actual relationships between the probable error and the average difference in consecutive positions have been calculated by the Pearson and Woodworth methods. The relationships are naturally negative, though not so much so as they might be, the figures being as follows:

TABLE VIII

RELATIONSHIP OF P. E. AND A. D. P.

Subject	W	P
A	57	–15
B	51	–30
C	75	–60
D	53	–33
E	64	–41

(Note that under W a figure above 50 indicates negative correlation.)

Let us consider more in detail the following portion of our results, positions 21–25 in the records of subject E. The grades here assigned to the cards in 21st–25th positions with their averages and m. v.'s are as follows:

TABLE IX

21	30	19	25	14	20	21.6	4.8
22	33	14	18	30	15	22.0	7.6
23	13	17	42	8	38	23.6	12.8
24	9	29	24	25	40	25.4	7.2
25	28	21	16	39	30	26.8	6.6

The weakness of the unsupported average and probable error as measures of the difference between two consecutive objects lies in the fact that they take no account of the coincidence of the grades which form them, and which ought to be a most important factor in the situation. Suppose, for example, we wish to determine E's attitude toward the cards whose averages place them 22d and 23d on the list. Out of the five judgments we see that in three cases, in two of a considerable margin, 22 was preferred over 21, and only the extreme fourth case gives it a slightly lower place. So much is not fully indicated in the m. v. The point is perhaps better illustrated in positions 24 and 25. In series I and IV there is extreme preference, outside the limits of the m. v.'s for 24 over 25, and the remainders show an almost equally

certain preference for the lower over the higher. The two cards need not really be close together at all; only now is one markedly preferred, now another. We can get a very different situation without altering average or m. v. in the slightest, Suppose the coincidence of these grades had been

24	9	29	24	25	40
25	16	30	21	28	39

We could then, perhaps, say with more confidence that 24 was preferred over 25. At any rate, the two results would have very different meanings, no difference appearing in the average or p. e., which are necessarily the same throughout.

In two consecutive positions from a series with much smaller probable errors the actual coincidence of the grades was as follows:

																						Av.	M. V.
A	.	.	.	8	5	5	9	6	2	2	1	5	2	3	1	4	6	2	7	4	4	5	7	4.5	1.7
B	.	.	.	7	3	4	2	3	6	7	4	4	5	10	5	10	5	5	6	3	6	7	1	5.1	1.2

There is .6 place difference in position and the p. e.'s of the averages do not overlap; yet in half the cases the lower position receives a higher grade than the higher. The grades cannot be rearranged so that this happens in more than twelve cases, they can be rearranged so that it happens in only three. The average and p. e. give no hint as to the nature of the coincidences, and their meaning is perhaps sufficient to warrant some special figure to express it.

Experiments on Color Vision

The apparatus used in these experiments[1] consisted of a series of 28 cards upon which were fixed, side by side, two silk skeins of differing colors. The colors were numbered 2, 4, 6, 8, etc., and the first card, known as 2–4, bore colors 2 and 4, the next 4 and 6, and so on up to 54–56, when the next bore 56 and then again the first color, 2. The colors thus ran through a complete circle, starting at the reds, and running through the yellows, greens, and purples back again to the reds. It was not attempted to have the series consist of saturated colors. The steps between the colors composing the pairs are not equal for sensation, and the original object of the experiment was to determine whether measurement by relative position would afford a means for stating the differences between the steps in a workable statistical form. Certain of the results are, however, germane to the present subject. The procedure was to have the subject arrange the pairs in order of the degree of their differences, the pair which differed least being counted 1, the next nearest as 2, the most dissimilar pair receiving a grade of 28. Arrangements were obtained from ten subjects, the order, positions, and mean variations being as follows:

[1] This material was being employed in a study of the quantitative measurement of color perception by Miss Mildred Focht of Columbia University, who kindly loaned it to me for the purpose of these experiments.

TABLE X

1	26–28	2.6	1.6
2	44–46	3.3	2.3
3	40–42	4.8	3.4
4	56–2	5.0	0.9
5	10–12	5.9	2.7
6	52–54	7.0	2.2
7	16–18	7.0	3.6
8	22–24	7.5	1.9
9	54–56	9.0	2.0
10	36–38	11.7	2.3
11	20–22	11.8	3.0
12	28–30	11.8	5.5
13	46–48	12.1	2.3
14	4–6	13.5	2.5
15	50–52	16.0	3.8
16	38–40	16.2	3.8
17	48–50	17.0	2.2
18	18–20	17.4	4.6
19	8–10	17.7	2.9
20	30–32	19.1	3.7
21	34–36	19.5	3.5
22	2–4	21.7	2.3
23	32–34	23.0	2.6
24	12–14	23.1	1.1
25	6–8	24.6	1.6
26	42–44	25.2	0.4
27	14–16	26.6	0.6
28	24–26	27.9	0.1

Color vision being something more objective than preference for souvenir postal cards, we find that the variability of the judgments is much smaller, the average m. v. of ten individuals for 50 postal cards being 10.8, and for estimations of the color differences but 2.4. The individual variations of each subject are distributed as follows:

TABLE XI

Subject	Variation														
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	3	5	6	7	3	..	..	1	..	1	2				
B	6	8	4	2	1	1	..	1	..	2					
C	6	6	9	2	..	2	1	1	..	..	..	..	1		
D	4	6	2	4	3	3	2	1	2	1					
E	4	7	5	7	1	2	2								
F	5	10	5	3	4	1									
G	8	8	4	2	4	1	1								
H	5	6	6	2	6	2	..	..	..	..	1				
I	5	7	6	6	..	..	2	1	..	..	1				
J	4	6	8	1	1	2	1	1	2	1	..	..	..	1	
Total	50	69	55	36	23	14	9	6	4	5	4		1	1	

There are six cases, two for A, one for C, and three for J, in which a pair is placed in a position differing from the average by more than three times the m. v. If such cases as these are not due

to chance, they demonstrate individual differences in color vision similar to those obtained in Henmon's experiments.[1] To make a rough determination of how far they might be due to chance, seven of the subjects arranged the series once more. These included subjects C and J, but it was unfortunately impossible to obtain another record from A. All of the divergences appear explicable as a result of chance. However, in calculating the m. v.'s of each subject in the two successive arrangements, the m. v.'s of each subject from his own judgment were considerably smaller than the mean of his variations from other subjects, the figures being as follows:

TABLE XII

Subject	C	D	E	G	H	I	J
Av. var. 2 succ. j. .	0.89	1.8	1.7	0.98	1.5	1.3	1.8
Av. var. j. 6 oth. ind.	2.3	3.3	2.3	1.7	2.4	2.3	3.2

There is still evidence of separate species in the judgments of each subject. The peculiar correspondence above noted between the amount of variation from one's own judgment and from the judgment of others appears here as in the postal cards. Between the two orders of Table XII there is 14 per cent of displacement; the more constant judges are the more accurate. As the objectivity of the experimental material increases, we should expect this correspondence to be closer.

[1] " The Time of Perception as a Measure of Differences in Sensation," *Archives of Phil., Psych., and Sci. Methods*, No. 8, 1906.

Experiments with Weights

It seemed best, for comparative purposes, to supplement the foregoing observations with a series of experiments in which the actual differences should be capable of determination by strictly objective methods. Weights are probably the most suitable material for this purpose. The apparatus consisted of six weights, 51, 53, 55, 57, 59, 61 grams, respectively.[1] The weights were made of dead black pasteboard boxes, $1\frac{3}{16}$ x $3\frac{1}{8}$ x $2\frac{1}{16}$ in., filled with lead and cotton to the required heaviness, and sealed. In the experiments the long axis of the weight was always toward the subject. The observations include 100 arrangements of the weights by one subject, and 10 arrangements by each of ten subjects. Of the subjects, G–J were normal individuals, the remainder being male patients in the hospital. Subject A is a man of 65, whose mental defect is a mixed paraphasia and object blindness. At the time of the experiments he could read and could name letters almost normally, but could not name objects, though they were recognized. Memory was much impaired. He co-operated conscientiously. B, æt. 52, is an early stage of general paresis, mildly euphoric. He co-operated willingly, but went at the test in a quick hit or miss fashion. C, æt. 72,

[1] The exact weights as measured on the scales of the physiological laboratory showed a practically constant excess of .4 gr. for each weight.

is a convalescent from a third attack of depression. Co-operated willingly, but showed a constant error in the shape of a tendency to leave the weights in the random order in which they were placed before him. D, æt. 64, manic-depressive, one previous attack of depression, at present mildly exhilarated. Co-operated willingly and conscientiously, but made frequent pauses between the arrangements on account of "fatigue." E, æt. 38, first attack of manic-depression, mixed phase, mildly exhilarated at time of experiment. Showed same tendency as C in leaving weights as at first placed. F, æt. 32, practical recovery from fourth attack of depression. Interested in experiment, and co-operated best of any of the patients, also doing the test exceptionally well. One other subject, a depression, actively lost interest after four trials, and failed to co-operate further. Each patient was held to a fixed system of procedure, analogous to that adopted by normal subjects. Only F would move the weights of his own accord, the others merely gave their judgments. The detail of their results *qua* from abnormal subjects I hope to discuss at some future time in connection with other observations. The data from the normal and abnormal subjects are quoted separately. As will be seen, two of the patients do normally, one exceptionally, well, while the remaining three do rather poorly. On the whole, there is nothing in the results to indicate a distinct species of per-

formance in the abnormal subject as a class. The general average is probably as valid for present purposes as one from ten normal subjects.

The following tables give the results of 100 arrangements by the single subject:

TABLE XIII

Series	Averages I	II	III	IV	V	VI	VII	VIII	IX	X	Av.	M.V.
61	1.3	1.2	1.8	1.5	1.4	1.0	1.2	1.3	2.0	1.1	1.4	.24
59	2.3	2.6	2.7	2.1	2.6	2.5	1.9	1.9	1.5	2.0	2.2	.32
57	3.4	3.3	2.9	2.7	2.4	2.8	3.0	2.9	3.2	3.1	3.0	.23
55	3.7	3.8	4.3	5.0	4.8	4.3	5.2	4.7	4.8	4.8	4.6	.40
53	5.5	5.4	4.1	4.5	4.8	5.1	4.1	5.0	4.5	4.6	4.8	.40
51	4.8	4.9	5.2	5.2	5.0	5.3	5.6	5.2	5.1	5.5	5.2	.28
Displacements of average	1	1	1	1	1	0	1	0	2	1	0.9	.29
Average of displacements	2.5	2.8	3.6	2.5	2.6	1.7	1.6	1.7	2.9	1.4	2.3	

Each column contains the average of a series of ten single arrangements. It will be noted that in only two cases out of the ten does the average order correspond with the objective one. Although the general average of the hundred arrangements gives the objective order, yet the displacements in the single series are hardly distributed according to chance. The fifth weight, 53, stands fifth with a position of 48 in the general average, but in five

series it stood above 55, in two below 51, and in only two cases did it stand in its proper position, thus accounting for seven out of the nine displacements of the averages of the series. In four of the seven cases, namely in series I, II, IV, and VII, the negative difference lies outside the limits of the probable error. VII is particularly striking on account of its high reliability throughout.

As the average should theoretically give the correct order no matter how poor the individual's judgment, the average of the displacements of each individual arrangement from the objective order is a better measure of difference between the accuracy of the successive series. The m. v. of the average order should also afford a measure of discriminativeness. According to both these measures the successive series show considerable practice, the average of the second five being a little over two-thirds that of the first five. The drop is unusually sudden. It may be observed that the displacement of the average and the average of displacements for the individual series are only moderately correlated. The average of displacements and the size of the m. v. are correlated within five displacements of their respective orders, or 11 per cent. We are here afforded an opportunity for examining empirically the accordance of an individual series with the average as a measure of the relative reliability of the successive series. As the average orders in the individual

series depart from the objective order, the method does not show up well. Between the accordance of each series of ten arrangements to their average, and the average of their displacements from the objective order, there are 20 displacements, 44 per cent; between the accordance of each series to their average, and the size of the m. v. in each series, there are 17 displacements, 38 per cent.[1] The mean variations of each series of ten arrangements from their averages (*i. e.*, the m. v.'s of the averages in the preceding table), are given below.

TABLE XIV

	Mean Variations										Av.
Series	I	II	III	IV	V	VI	VII	VIII	IX	X	
61	.58	.32	.96	.70	.64	.00	.32	.48	.60	.18	.48
59	.66	.92	1.10	.36	.92	.50	.36	.36	.60	.20	.60
57	1.18	.84	1.01	.82	.60	.80	.20	48	.98	.36	.73
55	.96	1.40	.82	.60	.80	.56	.48	.76	1.08	.79	.66
53	.60	.84	1.32	.80	.72	.76	.40	.60	.60	.80	.74
51	1.00	.74	.80	.80	.80	.66	.48	.80	.90	.60	.76
Av. m. v.	.83	.84	1.00	.68	.75	.55	.38	.58	.79	.49	.69

The average of the m. v.'s is naturally some-

[1] This is in part due to the fact that the poor judgments draw the end weights toward the middle while the good judgments keep them at the ends, thus getting a high variability for the extremes; if we take only the two middle weights, 57 and 55, we have from the average of displacements 17 displacements, or 38 per cent, instead of 44 per cent.

what larger than the m. v. of the averages, as given in Table XIII. It will be noted that the psychophysical relationship plays little part in these results; the difference between 51 and 53 should be greater than that between 59 and 61, but so far as can be judged from the results, 61 is more easily distinguished from 59 than 53 from 51. This is surprising, as the one hundred arrangements ought to be sufficient to bring out such a difference. The m. v.'s of the averages, as given in Table XIII, are smallest at the ends, as they arithmetically should be; but the averages of the m. v.'s, in Table XIV, seem to increase as the weights become smaller.

We may now compare the variation of the single subject through ten successive series, with the variation of ten different subjects through a single series of ten arrangements each. The results of these experiments are summarized in Tables XV and XVI.

The figures present the same general characteristics as those in Tables XIII and XIV. The single subject has varied from his own judgments a little less than the ten subjects among themselves, but this is in part due to practice, which brings down the m. v. If we take the m. v. of the first five series in which practice is not evident to any marked degree, and compare this with the variation of the four normal subjects, we see that the single subject has varied from himself rather more than the four normal subjects among themselves.

TABLE XV

Subject	Averages A	B	C	D	E	F	G	H	I	J	Av.	M. V.	Av. 6 path.	Av. 4 normal
wt. 61	1.4	1.7	1.7	1.5	2.4	1.5	1.3	1.8	1.1	1.4	1.6	0.26	1.7	1.4
59	2.8	2.8	4.0	2.1	2.7	2.3	2.3	2.6	1.9	2.0	2.5	0.43	2.6	2.2
57	2.8	3.7	4.0	3.0	2.6	2.6	3.4	3.0	3.0	3.1	3.1	0.32	3.1	3.1
55	4.1	4.1	3.4	4.3	4.3	4.2	3.7	4.4	4.1	4.6	4.1	0.24	4.0	4.2
53	4.3	3.8	2.8	5.1	4.3	4.9	5.5	5.0	5.0	4.3	4.5	0.60	4.2	4.8
51	5.6	4.9	4.9	5.0	4.8	5.6	4.8	5.2	5.8	5.0	5.2	0.32	5.2	5.2
Displacement of averages	0	1	4	1	1	0	1	0	0	1	0.9	. .	. . .	. . .
Av. of displacements	2.5	4.1	4.9	2.4	4.7	1.8	2.5	2.3	0.7	2.7	2.9	. .	3.4	2.1

TABLE XVI

Subject.	Mean Variations A	B	C	D	E	F	G	H	I	J	Av.	Av. 6 path.	Av. 4 normal
wt. 61	0.48	0.58	0.60	0.60	0.84	0.50	0.58	0.68	0.09	0.56	0.55	0.60	0.48
59	0.88	1.20	1.40	0.74	1.31	0.82	0.66	0.60	0.09	0.60	0.83	1.06	0.49
57	1.20	1.36	1.00	0.70	1.50	0.86	1.18	0.80	0.20	1.16	1.00	1.10	0.84
55	0.74	1.51	1.40	1.09	1.10	0.48	0.96	0.88	0.46	1.08	0.97	1.05	0.84
53	0.73	1.04	1.24	0.54	1.27	0.92	0.60	0.60	0.40	0.75	0.81	0.96	0.59
51	0.60	1.04	0.94	0.80	1.12	0.48	1.00	0.66	0.16	0.40	0.72	0.83	0.58
Av. m. v.	0.77	1.14	1.10	0.75	1.19	0.68	0.83	0.70	0.22	0.76	0.85	0.93	0.64

The figure for the single subject is .82, for the six patients it is .93, and for the four normal subjects .64. This is anomalous, for the variation of an individual should only approach the limit of the variability of the group and not exceed it.

Nevertheless, a striking contrast is formed to the relative variations in the repeated judgments of the postal cards, where each subject's judgments were a distinct species of their own.

In Table XV the record of subject C contains two very coarse deviations from the objective order. There is a remarkable overestimation of 53 and a lesser one of 55, while 57 and 59 have correspondingly low positions. It may be remembered that this subject showed a tendency to leave the weights as they were put before him, and in random arrangements 53 would ordinarily occupy a position higher than its objective one, 59 a lower. But so would 51 and 61, which are unaffected. Subject I underestimates 53, J overestimates it. Altogether, 53 is seen to have a very peculiar behavior.

Comparing, as in Tables XIII and XIV, the average of displacement with the average m. v., we find between them four displacements, 9 per cent. The order of discriminativeness of the ten subjects as measured by the accordance of their individual averages with the average of the ten, gives 14 displacements from the average of displacements and 15 from the size of the average m. v., 31 per cent and 33 per cent respectively. The displacements of the two middle weights, 57 and 55, from the average of displacements are 11, or 24 per cent instead of 31 per cent, for the whole six weights. This result thus agrees strikingly

with the result for the single subject. The final average order being correct in both cases, it would seem that, empirically, the number of displacements of an individual order from an average gives a better idea of its relative correctness than the precise arithmetical amount of its deviation from this order. It may then also be used in cases where there is no objective, only an average order, as in judgments of mental traits.

Evidence of the psychophysical relationship is again absent; 61 has a much smaller m. v. than 51, while those of 59 and 53 are practically equal. The m. v.'s are here largest in the middle, as they should be.

Conclusion

We have thus made a brief study of variability in three classes of judgment; first, the highly subjective feeling of preference for different sorts of pictures, second, the more objective judgment of color differences, and finally of a type of judgment whose accuracy could be readily measured by objective means. It has appeared that in the first class the judgments of each individual cluster about a mean which is true for that individual only, and which varies from that of any other individual more than twice as much as its own judgments vary from it; that in the second class, with the colors, the variability of the successive

judgments and those by different individuals markedly approached each other, but still preserved a significant difference; while in the third class, with the weights, we found that there might be even an excess of the individual variability over the "social." This comparison seems to afford, to a certain extent, a *quantitative criterion of the subjective.*

In objective fields those who vary least from their own judgments are, in the absence of constant error, those of the most reliable judgment; indeed, the constancy of our own opinions among themselves seems to be more important than their agreement with the standard of others. It is noteworthy that those who vary less from their own judgments are more likely to vary less from the judgments of others in the cards and colors than in the weights; it has been shown that this cannot be ascribed wholly to the small ranges with the weights.

It has again appeared in these experiments that even in those fields that we might ordinarily term most strictly objective, there are often certain relations between compared stimuli that are constant and peculiar to the individual. The same phenomenon appeared in Henmon's work on color-differentiation, two pairs of colors not necessarily standing in the same relation to each other with two individuals. The writer also observed it in experimenting with sounds of language, there oc-

curring a constant tendency to hear certain sounds rather than others, which differed with the individual. This is, however, most difficult to understand with our weights, for it would seem to indicate that the differences were not only of kind, but also of degree. The situation is not one that could be readily accounted for by displaced centres of gravity. This peculiar phenomenon, for which sensation habit is perhaps as good a term as any, is one that stands in much need of special and accurate investigation.